

AD \_\_\_\_\_

Award Number: DAMD17-98-1-8214

TITLE: Epigenetic Changes in DNA Methylation in Breast Cancer

PRINCIPAL INVESTIGATOR: Doctor Tim Huang

CONTRACTING ORGANIZATION: University of Missouri  
Columbia, Missouri 65211

REPORT DATE: September 2001

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

**REPORT DOCUMENTATION PAGE**Form Approved  
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

<b>1. AGENCY USE ONLY (Leave blank)</b>		<b>2. REPORT DATE</b> September 2001	<b>3. REPORT TYPE AND DATES COVERED</b> Final (1 Aug 98 - 1 Aug 01)	
<b>4. TITLE AND SUBTITLE</b> Epigenetic Changes in DNA Methylation in Breast Cancer			<b>5. FUNDING NUMBERS</b>  DAMD17-98-1-8214	
<b>6. AUTHOR(S)</b>  Doctor Tim Huang				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> University of Missouri Columbia, Missouri 65211  E-Mail: <a href="mailto:Huangh@health.missouri.edu">Huangh@health.missouri.edu</a>			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			<b>10. SPONSORING / MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b>  Report contains color.			<b>20020124 206</b>	
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for Public Release; Distribution Unlimited			<b>12b. DISTRIBUTION CODE</b>	
<b>13. ABSTRACT (Maximum 200 Words)</b> It is now clear that aberrant DNA methylation observed in cancer cells is not restricted to a few CpG islands, but affects multiple loci. When this epigenetic event occurs at the 5'-end of the regulatory region of genes, it is frequently associated with transcriptional silencing. To further investigate this widespread event in the tumor genome, we developed a novel microarray containing 7,776 short GC-rich tags tethered to glass slide surfaces. This DNA chip was used to study 17 paired tissues of breast tumors and normal controls. Amplicons, representing differential pools of methylated DNA fragments between tumors and normal controls, were co-hybridized to the microarray panel. Hypermethylation of multiple CpG island loci was then detected in a two-color fluorescence system. Approximately 1% (on average 83 loci) of these CpG islands examined were hypermethylated in this patient group. Hierarchical clustering segregated these tumors based on their methylation profiles and identified a group of CpG island loci that corresponds to the hormone-receptor status of breast cancer. This observation was independently confirmed by examining a single locus, the promoter of the human <i>glypican 3</i> gene, which was predominately hypermethylated in the hormone receptor-negative tumors. Our findings support the notion that hypermethylation of critical CpG island loci influences cancer development and produces distinct epigenetic signatures for particular tumor subtypes.				
<b>14. SUBJECT TERMS</b> Breast Cancer			<b>15. NUMBER OF PAGES</b> 31	
			<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> Unclassified	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> Unclassified	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> Unclassified	<b>20. LIMITATION OF ABSTRACT</b> Unlimited	

#### (4) Table of Contents

#### Page

Front cover-----	
Standard form (SF) 298, report documentation page-----	2
Table of content-----	3
Introduction-----	4
Body-----	4
Key research accomplishments-----	6
Reportable outcomes-----	7
Conclusions-----	7
References-----	7
Appendices-----	

## (5) Introduction

Cancer is generally thought to arise following multiple genetic mutations. More recently, a different type of alteration, i.e., CpG island hypermethylation, has been shown to be common in cancers (reviewed in Refs. 1-2). Unlike genetic mutations, this alteration is an epigenetic event that does not involve changes in nucleotide sequences. Nevertheless, accumulating evidence indicates that this heritable change has profound effects in tumorigenesis. When CpG island hypermethylation occurs at the 5'-end in the regulatory regions of genes (i.e., the promoter and the first exon), it may result in silencing of the corresponding genes (2). Both critical and non-critical CpG island loci could be affected and only those providing loss of key biochemical functions would render selective advantages for tumor cells. Thus far, CpG island hypermethylation has been shown to be an important mechanism for the transcriptional inactivation of more than 90 tumor-related genes in many types of cancers (see our web site: <http://www.missouri.edu/~hypermet>).

As a further step toward a comprehensive understanding of this epigenetic event in cancer, it is necessary to develop new techniques for genome-wide methylation analysis. Recently, we adapted an array-based strategy and developed DMH<sup>3</sup> (3) for high-throughput analysis of CpG island hypermethylation. Using a panel of CpG island tags arrayed on nylon membranes, DMH was successfully applied to detect specific methylation profiles in breast and ovarian tumors (4, 5).

To increase DMH throughput, we have made a complete transition from nylon membrane macroarrays to glass slide microarrays. A panel containing 7,776 CpG islands was generated and used to analyze 17 paired tissues of breast tumors and normal controls. Close to 6.5% (496 loci) of these tags exhibited hypermethylation at least once in the tumors analyzed. Hierarchical clustering classified breast tumors into groups based on their methylation profiles that correlated with clinically-related features.

## (6) Body

**a) Develop differential methylation hybridization (DMH) technique (corresponding to Task 1).** We previously developed the DMH technique for high-throughput screening of CpG island hypermethylation in cancer (3-5). To increase its throughput, we reconfigured DMH into a microarray-based system containing 7,776 short DNA tags prepared from a genomic library, CGI (6). Although the estimated 20,000-30,000 *MseI*-restricted CGI clones may not represent the entire repertoire of CpG islands in the genome, this library has provided a valuable resource for identifying aberrantly methylated loci in breast and ovarian cancer (4, 5). Approximately 25% of the 7,776 CpG island tags have so far been sequenced, and their sequence information has been deposited in GenBank.

For DMH screening, targets were prepared from a group of 17 paired breast tumors and normal samples. Unlike cDNA microarrays that use targets directly prepared from two different mRNA populations, targets generated for DMH involve a series of steps. One of the key steps requires the methylation-sensitive restriction of *MseI*-digested, linker-ligated DNA. In contrast to the previous use of a single endonuclease (3-5), in the present study DNA was digested sequentially with two methylation-sensitive endonucleases, *HpaII* and *BstUI*. This new approach enhances detection of CpG loci with extensive methylation in tumors, a situation more likely to result in chromatin condensation and subsequent silencing of corresponding genes (7). In addition, the approach reduces the risk of incomplete digestion when only one endonuclease is used for probing methylation differences. Low amplification cycles (20 cycles) were used for linker-PCR. This step prevented over-amplification of unrestricted repetitive sequences in the ligated DNA, and yet yielded sufficient PCR products for single or low-copy number CpG island loci. Genomic fragments containing aberrantly methylated sites were protected from the digestion and could be amplified by linker-PCR in the tumor sample, whereas the same fragments containing the unmethylated sites were cut and could not be amplified in the normal sample. The amplified products (or amplicons) therefore contained different pools of DNA fragments due to the differential methylation status of tumor relative to the control sample.

Fig. 1A shows representative DMH microarrays co-hybridized with fluorescently-labeled tumor and normal amplicons. CpG island tags whose signal intensities were slightly above the background or were devoid of hybridization signals represent the unmethylated loci in both tumor and normal samples; their genomic fragments were restricted away by the methylation-sensitive endonucleases prior to linker-PCR. "Yellow" spots

(Cy5/Cy3 = 1) represent equal amounts of bound DNA from each amplicon, indicating no methylation differences between tumor and normal genomes. In some instances, these “yellow” spots represented CpG island tags that do not contain the internal *HpaII* or *BstUI* recognition sites and have equal copy numbers in both tumor and control DNA. CpG island tags hybridized predominately with the tumor amplicon, but not with the normal amplicon, appear as “red” spots. We set an arbitrary cut-off of 1.5 for the Cy5/Cy3 ratio – i.e., loci with ratios  $\geq 1.5$  were identified as hypermethylated in tumors. It should be noted that the magnitude of the Cy5/Cy3 ratio does not necessarily reflect the extent of hypermethylation, as the target preparation is PCR-based. An example is shown in Fig. 1B,C,D – The SC76F1 locus appeared to be hypermethylated in the breast tumor of patient 241, but not in patients 151 and 109. Though less frequently, we also encountered “green” spots (Cy5/Cy3  $\leq 0.5$ ) by DMH, denoting the presence of hypomethylated sequences in the tumor genome (see Fig. 1C). Sequence analysis indicated that most of the “green” spots in the microarray panel are repetitive elements, which are often methylated in non-tumor cells (8). Hypomethylation is observed in heterochromatin or  $\alpha$ -satellite DNA in cancer cells, but has not commonly been seen in single-copy CpG islands (8).

We conducted a confirmation study to determine whether the cut-off ratio ( $\geq 1.5$ ) could accurately identify hypermethylated loci. This was performed by Southern analysis using amplicons, originally prepared for DMH, as hybridization templates (Fig. 2). For example, the SC76F1 probe detected a DNA fragment (770-bp) in the tumor amplicons 65, 119, 167, and 241, but not in their paired-normal amplicons. The presence of this amplified fragment in tumor amplicons was because methylated CpG sites within the SC76F1 locus were insensitive to restriction by endonucleases. This 770-bp fragment was not detected in tumor amplicons 83 and 151 or in their controls, indicating this locus was not hypermethylated in these tumors. In some cases (e.g., CpG5H12, SC13B2, and CpG15E5), normal breast samples had detectable methylation, but methylation of the loci was more extensive in their tumors. It is possible that this low level of pre-existing methylation is present in normal samples (3, 4). The results of 7 positive loci based on this Southern approach as well as their Cy5/Cy3 ratios derived from DMH are summarized in Table 1. The Southern findings appeared to be consistent with DMH results in the majority (95%) of the loci analyzed. The inconsistency between these methods is, in part, due to the imprecise nature of an overall Cy5/Cy3 normalization factor for a given DMH image. As a result, the cut-off ratio set for the identification of DNA hypermethylation may be too high to exclude some positive loci and too low to include some unmethylated loci in a given tumor. One remedy to this problem is to perform sufficient Southern amplicon analyses and then readjust the normalization factor to clearly define the hypermethylation boundary for most of the CpG island loci. As with other microarray approaches, we expect low levels of false-positive/negative findings in DMH. This should only have a minimal impact on the subsequent data analysis due to a large number of hypermethylated loci identified by DMH.

Of the 7,776 CpG island loci probed by our DMH amplicons, loci having a normalized Cy5/Cy3 ratio  $\geq 1.5$  were selected for cluster analyses. Across the 17 tumors studied, the average number of hypermethylated loci was 83 (ranging from 15 to 207), consistent with the notion that hypermethylation is infrequent and occurs in ~1% of CpG islands in the breast tumor genome (9). As tumors exhibited hypermethylation in some shared loci while having other uniquely methylated CpG islands, the total number of loci identified in this patient group was 496. The Stanford hierarchical algorithm (10) was then used to extract maximum information from these positive loci for tumor classification (Fig. 3A). Loci with little or no methylation differences between tumor and normal DNA were left blank, while hypermethylated loci in tumors were colored “red”. The resulting dendrogram divided 15 of these 17 tumors into three clusters, which also were confirmed using a separate clustering algorithm, the multidimensional 3-D scatterplot (Fig. 3B). Sections of the dendrogram that contribute significantly to segregating these clusters are shown in Fig. 3C. Two of these clusters were made up of tumors largely with negative ER/PR<sup>3</sup> status, a group commonly less responsive to certain forms of chemotherapy; CpG island loci in block 2 were frequently methylated in these two patient groups. Block 1 represented loci predominately methylated in one of these two groups (patients #241, 31, 243, 91, 159, and 155), while loci in block 3 were methylated more frequently in the other group (patients #45, 43, 65, 167, and 119). The remaining cluster, comprised of ER/PR-positive tumors (patients #157, 151, 129 and 109), was largely devoid of methylation in these CpG island loci.

**b) Characterization of hypermethylated CpG island loci in breast cancer cell lines and primary tumors (corresponding to Tasks 2 & 3).** This methylation profile analysis has led to the identification of CpG

island clusters that can distinguish ER/PR-negative from ER/PR-positive tumors. To determine whether a single CpG island locus within these clusters contributes to this discrimination, we examined SC76F1 in greater detail. The SC76F1 sequence matched a CpG island region located within the promoter of *GPC3* gene at chromosome Xq26 (11). This gene encodes a glypican integral membrane protein and is mutated in the Simpson-Golabi-Behmel syndrome, an X-linked condition involving multiple embryonic neoplasms and pre- and postnatal overgrowth (12). Interestingly, hypermethylation of this gene has recently been reported in several ovarian cancer cell lines (13). We determined the methylation status of *GPC3* promoter in 45 breast tumors and 10 normal breast tissue samples by Southern analysis (Fig. 4A). Both methylated (490-bp) and unmethylated (130 and 360-bp) alleles of *GPC3* were observed in the normal breast samples analyzed, consistent with an X-linked status of this gene (11). Various degrees of hypermethylation were detected in 38 of 45 (84%) breast tumors, while the remaining 7 tumors showed no detectable methylation changes. This result also confirms the DMH microarray findings. The exception is two tumors (#151 and 129) in which a diminutive level of methylation was detected by Southern analysis but not by DMH. Statistical analysis (*t* test) reveals a significant difference ( $P=0.005$ ) in *GPC3* promoter hypermethylation between the hormone receptor-positive tumors and the hormone receptor-negative tumors. We further segregated these tumors according to the levels of *GPC3* hypermethylation (Fig. 4B). It is apparent that high levels of *GPC3* methylation are predominantly associated with the hormone receptor-negative group, consistent with the earlier finding based on the cluster analysis. Hypermethylation of the *GPC3* promoter resulted in loss of its expression *in vitro*, suggesting a functional consequence of this methylation-silenced gene in breast tumor growth (Fig. 4C,D).

The present study is the first to provide such a detailed analysis of the complex methylation alterations in the human breast tumor genome. This type of microarray-based analysis is expected to move this field into the epigenomics era. It is now clear that aberrant DNA methylation has a causative role, rather than a bystander effect, in the development of cancer (1, 2). Many promoter CpG island sequences are susceptible to this epigenetic alteration. As a result, the expression of genes that govern key functions of the cell may become silent, leading to clonal proliferation of tumor cells. Differential susceptibility of critical CpG island loci to DNA hypermethylation may therefore influence the development of different tumor types and produces unique molecular signatures that are associated with clinicopathological characteristics of cancer patients. Dissecting these complex epigenetic profiles can also lead to the identification of tumor-suppressors that are silenced via DNA hypermethylation in particular tumor subtypes. Thus, this analysis has diagnostic potential and provides patient-specific methylation profiles that may predict those responsive to demethylation treatment (14). Gene expression profiles produced by cDNA microarrays have already been used to differentiate different types of lymphoma (15), leukemia (16), melanoma (17), and breast cancer (18). In this regard, DMH not only provides further information regarding potential mechanisms underlying gene expression, but also offers an alternative to cDNA microarrays for molecular classification of tumors. The cDNA microarray approach requires targets derived from mRNA, which is more labile and difficult to isolate directly from clinical specimens. By contrast, our CpG island microarray uses DNA targets, which is more stable and easier to isolate from patients' specimens.

In summary, we have presented data from our newly developed DMH microarray method that illustrates its utility both in the basic understanding of CpG island hypermethylation in cancer and also in a translational role as a potential method for molecular classification of tumors and prediction of responsiveness to demethylating agents. Studies such as this may open up a whole new era of individually designed pharmacologic agents for each patient based on solid molecular evidence of efficacy.

## (7) Key Research Accomplishments

- A panel of 7,776 CpG island tags has been established for genome-wide methylation analysis using DMH microarray.
- Microarray analysis identified a group of CpG island loci that corresponds to the hormone-receptor status of breast cancer.
- The promoter of the human *glypican 3* gene, which was predominately hypermethylated in the hormone receptor-negative tumors.

## (8) Reportable Outcomes

- One manuscript has been submitted for publication in *Cancer Research*. The same manuscript was presented in the 2001 AACR meeting in New Orleans (see the Appendix).

## (9) Conclusion

In summary, we have presented data from our newly developed DMH microarray method that illustrates its utility both in the basic understanding of CpG island hypermethylation in cancer and also in a translational role as a potential method for molecular classification of tumors and prediction of responsiveness to demethylating agents. Studies such as this may open up a whole new era of individually designed pharmacologic agents for each patient based on solid molecular evidence of efficacy.

## (10) References

1. Jones, P. A. and Laird, P. W. Cancer epigenetics comes of age. *Nat. Genet.* 21: 163-167, 1999.
2. Baylin, S. B. and Herman, J. G. DNA hypermethylation in tumorigenesis: Epigenetics joins genetics. *Trends Genet.* 16: 168-174, 2000.
3. Huang, T. H.-M., Perry, M. R., and Laux, D. E. Methylation profiling of CpG islands in human breast cancer cells. *Hum. Mol. Genet.* 8: 459-470, 1999.
4. Yan, P. S., Perry, M. R., Laux, D. E., Asare, A., Caldwell, C. W., and Huang, T. H.-M. CpG island arrays: An application toward deciphering epigenetic signatures of breast cancer. *Clin. Cancer Res.* 6: 1432-1438, 2000.
5. Ahluwalia, A., Yan, P. S., Bigsby, R. M., Hurteau, J. A., Jung, S. H., Huang, T. H.-M., and Nephew, K. P. Analysis of CpG island hypermethylation in human ovarian cancer using differential methylation hybridization. *Gyn. Onc.* (in press).
6. Cross, S. H., Charlton, J. A., Nan, X., and Bird, A. P. Purification of CpG islands using a methylated DNA binding column. *Nat. Genet.* 6: 236-244, 1994.
7. Baylin, S. B., Esteller, M., Rountree, M. R., Bachman, K. E., Schuebel, K., and Herman, J. G. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum. Mol. Genet.* 10: 687-692, 2001.
8. Ehrlich, M. DNA hypomethylation and cancer. *In: M. Ehrlich (ed.) DNA Alterations in Cancer*, pp. 273-291. Natick, MA: Eaton Publishing, 2000.
9. Costello, J. F., Fruhwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., Wright, F. A., Feramisco, J. D., Peltomaki, P., Lang, J. C., Schuller, D. E., Yu, L., Bloomfield, C. D., Caligiuri, M. A., Yates, A., Nishikawa, R., Su Huang, H., Petrelli, N. J., Zhang, X., O'Dorisio, M. S., Held, W. A., Cavenee, W. K., and Plass, C. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat. Genet.* 24: 132-138, 2000.
10. Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA.* 95: 14863-14868, 1998.
11. Huber, R., Hansen, R. S., Strazzullo, M., Pengue, G., Mazzarella, R., D'Urso, M., Schlessinger, D., Pilia, G., Gartler, S. M., and D'Esposito, M. DNA methylation in transcriptional repression of two differentially expressed X-linked genes, GPC3 and SYBL1. *Proc. Natl. Acad. Sci. USA.* 96: 616-621, 1999.
12. Veugelers, M., Cat, B. D., Muyldermans, S. Y., Reekmans, G., Delande, N., Frints, S., Legius, E., Fryns, J. P., Schrandt-Stumpel, C., Weidle, B., Magdalena, N., and David, G. Mutational analysis of the GPC3/GPC4 glypican gene cluster on Xq26 in patients with Simpson-Golabi-Behmel syndrome: identification of loss-of-function mutations in the GPC3 gene. *Hum. Mol. Genet.* 9: 1321-1328, 2000.
13. Lin, H., Huber, R., Schlessinger, D., and Morin, P. J. Frequent Silencing of the GPC3 Gene in Ovarian Cancer Cell Lines. *Cancer Res.* 59: 807-810, 1999.

14. Santini, V., Kantarjian, H. M., and Issa, J. P. Changes in DNA methylation in neoplasia: Pathophysiology and therapeutic implications. *Ann. Intl. Med.* 134: 573-586, 2001.
15. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., and Staudt, L. M. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403: 503-511, 2000.
16. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286: 531-537, 1999.
17. Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., Sampas, N., Dougherty, E., Wang, E., Marincola, F., Gooden, C., Lueders, J., Glatfelter, A., Pollock, P., Carpten, J., Gillanders, E., Leja, D., Dietrich, K., Beaudry, C., Berens, M., Alberts, D., and Sondak, V. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406: 536-540, 2000.
18. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A. L., Brown, P. O., and Botstein, D. Molecular portraits of human breast tumours. *Nature* 406: 747-752, 2000.



# **Dissecting Complex Epigenetic Alterations in Breast Cancer Using CpG**

## **Island Microarrays**

Pearlly S. Yan, Chuan-Mu Chen, Huidong Shi, Farahnaz Rahmatpanah, Susan H. Wei,  
Charles W. Caldwell, and Tim Hui-Ming Huang<sup>1,2</sup>

Departments of Pathology and Anatomical Sciences, Ellis Fischel Cancer Center, University of  
Missouri School of Medicine, Columbia, Missouri 65203 [*P. S. Y., H.S., F. R., S. H. W., C. W.*  
*C., T. H.-M. H.*]; Department of Zoology, Life Science College, National Chung Hsing  
University, Taiwan, Republic of China [*C.-M. C.*]

**Running Title:** DNA Hypermethylation in Breast Cancer

**Keywords:** Breast Cancer, CpG island, Microarray, DNA Methylation

<sup>1</sup>This work was supported by National Cancer Institute grants CA-69065 and CA-84701 and by  
U. S. Army Medical Research Command grant DAMD17-98-1-8214. C.-M. C was a visiting  
fellow supported by National Science Council, Taiwan (NSC39073F). S. H. W. was supported  
by a postdoctoral fellowship from Cancer Research Center, Inc.

<sup>2</sup>To whom requests for reprint should be addressed, at the Department of Pathology and  
Anatomical Sciences, Ellis Fischel Cancer Center, University of Missouri, 115 Business Loop I-  
70 West, Columbia, MO 65203, USA. Tel: 573-882-1276; Fax 573-884-5206;  
E-mail: huangh@health.missouri.edu

<sup>3</sup>The abbreviations used are; DMH, differential methylation hybridization; aa-dUTP, amino-allyl  
dUTP; *GPC3*, *glypican 3*; ER/PR, estrogen- and progesterone-receptors.

## Abstract

It is now clear that aberrant DNA methylation observed in cancer cells is not restricted to a few CpG islands, but affects multiple loci. When this epigenetic event occurs at the 5'-end of the regulatory region of genes, it is frequently associated with transcriptional silencing. To further investigate this widespread event in the tumor genome, we developed a novel microarray containing 7,776 short GC-rich tags tethered to glass slide surfaces. This DNA chip was used to study 17 paired tissues of breast tumors and normal controls. Amplicons, representing differential pools of methylated DNA fragments between tumors and normal controls, were co-hybridized to the microarray panel. Hypermethylation of multiple CpG island loci was then detected in a two-color fluorescence system. Approximately 1% (on average 83 loci) of these CpG islands examined were hypermethylated in this patient group. Hierarchical clustering segregated these tumors based on their methylation profiles and identified a group of CpG island loci that corresponds to the hormone-receptor status of breast cancer. This observation was independently confirmed by examining a single locus, the promoter of the human *glypican 3* gene, which was predominately hypermethylated in the hormone receptor-negative tumors. Our findings support the notion that hypermethylation of critical CpG island loci influences cancer development and produces distinct epigenetic signatures for particular tumor subtypes.

## Introduction

Cancer is generally thought to arise following multiple genetic mutations. More recently, a different type of alteration, i.e., CpG island hypermethylation, has been shown to be common in cancers (reviewed in Refs.1-2). Unlike genetic mutations, this alteration is an epigenetic event that does not involve changes in nucleotide sequences. Nevertheless, accumulating evidence indicates that this heritable change has profound effects in tumorigenesis. When CpG island hypermethylation occurs at the 5'-end in the regulatory regions of genes (i.e., the promoter and the first exon), it may result in silencing of the corresponding genes (2). Both critical and non-critical CpG island loci could be affected and only those providing loss of key biochemical functions would render selective advantages for tumor cells. Thus far, CpG island hypermethylation has been shown to be an important mechanism for the transcriptional inactivation of more than 90 tumor-related genes in many types of cancers (see our web site: <http://www.missouri.edu/~hypermet>).

As a further step toward a comprehensive understanding of this epigenetic event in cancer, it is necessary to develop new techniques for genome-wide methylation analysis. Recently, we adapted an array-based strategy and developed DMH<sup>3</sup> (3) for high-throughput analysis of CpG island hypermethylation. Using a panel of CpG island tags arrayed on nylon membranes, DMH was successfully applied to detect specific methylation profiles in breast and ovarian tumors (4, 5).

To increase DMH throughput, we have made a complete transition from nylon membrane macroarrays to glass slide microarrays. A panel containing 7,776 CpG islands was generated and used to analyze 17 paired tissues of breast tumors and normal controls. Close to 6.5% (496 loci) of these tags exhibited hypermethylation at least once in the tumors analyzed. Hierarchical

clustering classified breast tumors into groups based on their methylation profiles that correlated with clinically-related features.

## **Materials and Methods**

**Tissue Samples and Cell Lines.** Tumor tissues were obtained from patients undergoing mastectomy prior to chemotherapy at Ellis Fischel Cancer Center, Columbia, MO, in compliance with our Institutional Review Board. Adjacent tumor-free parenchyma 3-5 cm away from the tumor area was obtained from each patient to serve as a paired control. Breast cancer cell lines T47-D, ZR-75, Hs578t, MCF-7, and MDA-MB-231 and MDA-MB-468 were acquired from various resources as described (3) and routinely maintained in our laboratory. High-molecular-weight genomic DNA was isolated using the QIAamp Tissue Kit<sup>TM</sup> (Qiagen).

**Preparation of CpG Island Microarray.** A total of 7,776 CpG island clones were individually organized in 96-well culture chambers as master plates. This included 10 preselected *Mse*I-tags that act as positive controls because they are known to lack the test methylation-sensitive sites. A fraction (~1  $\mu$ l) of each clone was transferred to a well of separate 96-well PCR tubes using the MULTI-PRINT<sup>TM</sup> replicator. CpG island inserts (0.2 to 2-kb) from these clones were amplified by PCR as described (3). The primers immediately flanking the inserts are HGMP 3558: 5'-CGG CCG CCT GCA GGT CTG ACC ATA A and HGMP 3559: 5'-AAC GCG TTG GGA GCT CTC CCA TAA (6). To ensure the reproducibility of each PCR and to prevent cross-contamination among bacterial clones in microplates, amplified inserts were individually verified using a 96-well format gel electrophoresis system (Cascade Biologics). The Affymetrix/GMS 417 microarrayer arrayed unpurified PCR products (~0.02  $\mu$ l per dot, 0.1

µg/µl), in the presence of 20% DMSO, as microdots (150 µm diameter spaced at 300 µm) on poly-L-lysine-coated microscope slides. Spotted DNA was denatured before use.

**Amplicon Generation.** The *MseI* digestion was performed in a 20 µl volume/sample containing ~2 µg genomic DNA. The digests were purified and their sticky ends ligated with 0.5 nmol of unphosphorylated linkers H-24/H-12 as described (3). The oligonucleotide sequences were as follows: H-24: 5'-AGG CAA CTG TGC TAT CCG AGG GAT and H-12: 5'-TAA TCC CTC GGA. The ligated DNA was digested with methylation-sensitive endonucleases *Bst*UI and *Hpa*II (New England Biolabs). PCR reactions were performed using the digests as templates and subjected to 20 cycles of amplification. The amplified products were purified, a portion of the products was reserved for Southern analysis, and the rest was used for fluorescence labeling.

**Microarray Hybridization and Data Analysis.** Incorporation of aa-dUTP<sup>3</sup> into amplicons (5 µg) was conducted using the Multiprime DNA labeling system (Life Technologies). Cy5 and Cy3 fluorescent dyes were coupled to aa-dUTP-labeled tumor and normal amplicons, respectively, and co-hybridized to the microarray panel. Hybridization and the post-hybridization washing protocols were according to DeRisi *et al.* (<http://www.microarrays.org>). Hybridized slides were scanned with the GenePix 4000A scanner (Axon) and the acquired images were analyzed with the software GenePix Pro3.0. Because Cy5 and Cy3 labeling efficiencies varied among samples, we determined a global normalization factor for each microarray image. The effectiveness of the normalization factor was evaluated using 10 internal positive controls, in which their adjusted Cy5/Cy3 ratios were expected to be 1. The adjusted ratio for each CpG island locus was calculated and the data generated were analyzed using the complete-linkage hierarchical clustering algorithm (<http://rana.stanford.edu/clustering>). All ratios were log-transformed prior to analysis and the resultant dendrogram linked related CpG island loci into a

phylogenetic tree whose branch lengths represented the degree of similarity between sets of hypermethylated loci. A nonhierarchical clustering algorithm (Partek Fuzzy C-Means protocol) and a statistical/visualization procedure (Multidimensional Scaling) was also used to analyze the methylation data set. This completely unstructured approach analyzed the methylation data in an entirely local fashion and produced an unorganized collection of clusters projected in a 3-dimensional view.

**Southern Analysis.** One  $\mu\text{g}$  PCR products (or amplicons) or 10  $\mu\text{g}$  genomic DNA digested with *MseI* and/or *BstUI* were separated on 1.0% agarose gels and transferred to nylon membranes. Short fragments (200 to 250-bp) from selected CpG island clone inserts were PCR-amplified and used as probes for Southern hybridization as described (3).

**Northern and RT-PCR Analyses.** Total cellular RNA was isolated from cells using the RNeasy Total RNA System (Qiagen). Ten  $\mu\text{g}$  of RNA was electrophoresed on a 1.5 % agarose gel and subjected to northern analysis using a *GPC3*<sup>3</sup> cDNA probe. RT-PCR was conducted using primers 5'-ATC CTG TAT ACC TCC TCC AG and 5'-ATC CAT GCA AAG AGA GAA CG for *GPC3* cDNA. The levels of *GPC3* mRNA were normalized with the level of  *$\beta$ -actin* mRNA as described in an earlier study (3).

## Results and Discussion

We previously developed the DMH technique for high-throughput screening of CpG island hypermethylation in cancer (3-5). To increase its throughput, we reconfigured DMH into a microarray-based system containing 7,776 short DNA tags prepared from a genomic library, CGI (6). Although the estimated 20,000-30,000 *MseI*-restricted CGI clones may not represent the entire repertoire of CpG islands in the genome, this library has provided a valuable resource for

identifying aberrantly methylated loci in breast and ovarian cancer (4, 5). Approximately 25% of the 7,776 CpG island tags have so far been sequenced, and their sequence information has been deposited in GenBank.

For DMH screening, targets were prepared from a group of 17 paired breast tumors and normal samples. Unlike cDNA microarrays that use targets directly prepared from two different mRNA populations, targets generated for DMH involve a series of steps. One of the key steps requires the methylation-sensitive restriction of *MseI*-digested, linker-ligated DNA. In contrast to the previous use of a single endonuclease (3-5), in the present study DNA was digested sequentially with two methylation-sensitive endonucleases, *HpaII* and *BstUI*. This new approach enhances detection of CpG loci with extensive methylation in tumors, a situation more likely to result in chromatin condensation and subsequent silencing of corresponding genes (7). In addition, the approach reduces the risk of incomplete digestion when only one endonuclease is used for probing methylation differences. Low amplification cycles (20 cycles) were used for linker-PCR. This step prevented over-amplification of unrestricted repetitive sequences in the ligated DNA, and yet yielded sufficient PCR products for single or low-copy number CpG island loci. Genomic fragments containing aberrantly methylated sites were protected from the digestion and could be amplified by linker-PCR in the tumor sample, whereas the same fragments containing the unmethylated sites were cut and could not be amplified in the normal sample. The amplified products (or amplicons) therefore contained different pools of DNA fragments due to the differential methylation status of tumor relative to the control sample.

Fig. 1A shows representative DMH microarrays co-hybridized with fluorescently-labeled tumor and normal amplicons. CpG island tags whose signal intensities were slightly above the background or were devoid of hybridization signals represent the unmethylated loci in both

tumor and normal samples; their genomic fragments were restricted away by the methylation-sensitive endonucleases prior to linker-PCR. “Yellow” spots ( $Cy5/Cy3 = 1$ ) represent equal amounts of bound DNA from each amplicon, indicating no methylation differences between tumor and normal genomes. In some instances, these “yellow” spots represented CpG island tags that do not contain the internal *HpaII* or *BstUI* recognition sites and have equal copy numbers in both tumor and control DNA. CpG island tags hybridized predominately with the tumor amplicon, but not with the normal amplicon, appear as “red” spots. We set an arbitrary cut-off of 1.5 for the  $Cy5/Cy3$  ratio – i.e., loci with ratios  $\geq 1.5$  were identified as hypermethylated in tumors. It should be noted that the magnitude of the  $Cy5/Cy3$  ratio does not necessarily reflect the extent of hypermethylation, as the target preparation is PCR-based. An example is shown in Fig. 1B,C,D – The SC76F1 locus appeared to be hypermethylated in the breast tumor of patient 241, but not in patients 151 and 109. Though less frequently, we also encountered “green” spots ( $Cy5/Cy3 \leq 0.5$ ) by DMH, denoting the presence of hypomethylated sequences in the tumor genome (see Fig. 1C). Sequence analysis indicated that most of the “green” spots in the microarray panel are repetitive elements, which are often methylated in non-tumor cells (8). Hypomethylation is observed in heterochromatin or  $\alpha$ -satellite DNA in cancer cells, but has not commonly been seen in single-copy CpG islands (8).

We conducted a confirmation study to determine whether the cut-off ratio ( $\geq 1.5$ ) could accurately identify hypermethylated loci. This was performed by Southern analysis using amplicons, originally prepared for DMH, as hybridization templates (Fig. 2). For example, the SC76F1 probe detected a DNA fragment (770-bp) in the tumor amplicons 65, 119, 167, and 241, but not in their paired-normal amplicons. The presence of this amplified fragment in tumor amplicons was because methylated CpG sites within the SC76F1 locus were insensitive to



restriction by endonucleases. This 770-bp fragment was not detected in tumor amplicons 83 and 151 or in their controls, indicating this locus was not hypermethylated in these tumors. In some cases (e.g., CpG5H12, SC13B2, and CpG15E5), normal breast samples had detectable methylation, but methylation of the loci was more extensive in their tumors. It is possible that this low level of pre-existing methylation is present in normal samples (3, 4). The results of 7 positive loci based on this Southern approach as well as their Cy5/Cy3 ratios derived from DMH are summarized in Table 1. The Southern findings appeared to be consistent with DMH results in the majority (95%) of the loci analyzed. The inconsistency between these methods is, in part, due to the imprecise nature of an overall Cy5/Cy3 normalization factor for a given DMH image. As a result, the cut-off ratio set for the identification of DNA hypermethylation may be too high to exclude some positive loci and too low to include some unmethylated loci in a given tumor. One remedy to this problem is to perform sufficient Southern amplicon analyses and then readjust the normalization factor to clearly define the hypermethylation boundary for most of the CpG island loci. As with other microarray approaches, we expect low levels of false-positive/negative findings in DMH. This should only have a minimal impact on the subsequent data analysis due to a large number of hypermethylated loci identified by DMH.

Of the 7,776 CpG island loci probed by our DMH amplicons, loci having a normalized Cy5/Cy3 ratio  $\geq 1.5$  were selected for cluster analyses. Across the 17 tumors studied, the average number of hypermethylated loci was 83 (ranging from 15 to 207), consistent with the notion that hypermethylation is infrequent and occurs in  $\sim 1\%$  of CpG islands in the breast tumor genome (9). As tumors exhibited hypermethylation in some shared loci while having other uniquely methylated CpG islands, the total number of loci identified in this patient group was 496. The Stanford hierarchical algorithm (10) was then used to extract maximum information from these

positive loci for tumor classification (Fig. 3A). Loci with little or no methylation differences between tumor and normal DNA were left blank, while hypermethylated loci in tumors were colored “red”. The resulting dendrogram divided 15 of these 17 tumors into three clusters, which also were confirmed using a separate clustering algorithm, the multidimensional 3-D scatterplot (Fig. 3B). Sections of the dendrogram that contribute significantly to segregating these clusters are shown in Fig. 3C. Two of these clusters were made up of tumors largely with negative ER/PR<sup>3</sup> status, a group commonly less responsive to certain forms of chemotherapy; CpG island loci in block 2 were frequently methylated in these two patient groups. Block 1 represented loci predominately methylated in one of these two groups (patients #241, 31, 243, 91, 159, and 155), while loci in block 3 were methylated more frequently in the other group (patients #45, 43, 65, 167, and 119). The remaining cluster, comprised of ER/PR-positive tumors (patients #157, 151, 129 and 109), was largely devoid of methylation in these CpG island loci.

This methylation profile analysis has led to the identification of CpG island clusters that can distinguish ER/PR-negative from ER/PR-positive tumors. To determine whether a single CpG island locus within these clusters contributes to this discrimination, we examined SC76F1 in greater detail. The SC76F1 sequence matched a CpG island region located within the promoter of *GPC3* gene at chromosome Xq26 (11). This gene encodes a glypican integral membrane protein and is mutated in the Simpson-Golabi-Behmel syndrome, an X-linked condition involving multiple embryonic neoplasms and pre- and postnatal overgrowth (12). Interestingly, hypermethylation of this gene has recently been reported in several ovarian cancer cell lines (13). We determined the methylation status of *GPC3* promoter in 45 breast tumors and 10 normal breast tissue samples by Southern analysis (Fig. 4A). Both methylated (490-bp) and unmethylated (130 and 360-bp) alleles of *GPC3* were observed in the normal breast samples

analyzed, consistent with an X-linked status of this gene (11). Various degrees of hypermethylation were detected in 38 of 45 (84%) breast tumors, while the remaining 7 tumors showed no detectable methylation changes. This result also confirms the DMH microarray findings. The exception is two tumors (#151 and 129) in which a diminutive level of methylation was detected by Southern analysis but not by DMH. Statistical analysis (*t* test) reveals a significant difference ( $P=0.005$ ) in *GPC3* promoter hypermethylation between the hormone receptor-positive tumors and the hormone receptor-negative tumors. We further segregated these tumors according to the levels of *GPC3* hypermethylation (Fig. 4B). It is apparent that high levels of *GPC3* methylation are predominantly associated with the hormone receptor-negative group, consistent with the earlier finding based on the cluster analysis. Hypermethylation of the *GPC3* promoter resulted in loss of its expression *in vitro*, suggesting a functional consequence of this methylation-silenced gene in breast tumor growth (Fig. 4C,D).

The present study is the first to provide such a detailed analysis of the complex methylation alterations in the human breast tumor genome. This type of microarray-based analysis is expected to move this field into the epigenomics era. It is now clear that aberrant DNA methylation has a causative role, rather than a bystander effect, in the development of cancer (1, 2). Many promoter CpG island sequences are susceptible to this epigenetic alteration. As a result, the expression of genes that govern key functions of the cell may become silent, leading to clonal proliferation of tumor cells. Differential susceptibility of critical CpG island loci to DNA hypermethylation may therefore influence the development of different tumor types and produces unique molecular signatures that are associated with clinicopathological characteristics of cancer patients. Dissecting these complex epigenetic profiles can also lead to the identification of tumor-suppressors that are silenced via DNA hypermethylation in particular tumor subtypes.

Thus, this analysis has diagnostic potential and provides patient-specific methylation profiles that may predict those responsive to demethylation treatment (14). Gene expression profiles produced by cDNA microarrays have already been used to differentiate different types of lymphoma (15), leukemia (16), melanoma (17), and breast cancer (18). In this regard, DMH not only provides further information regarding potential mechanisms underlying gene expression, but also offers an alternative to cDNA microarrays for molecular classification of tumors. The cDNA microarray approach requires targets derived from mRNA, which is more labile and difficult to isolate directly from clinical specimens. By contrast, our CpG island microarray uses DNA targets, which is more stable and easier to isolate from patients' specimens.

In summary, we have presented data from our newly developed DMH microarray method that illustrates its utility both in the basic understanding of CpG island hypermethylation in cancer and also in a translational role as a potential method for molecular classification of tumors and prediction of responsiveness to demethylating agents. Studies such as this may open up a whole new era of individually designed pharmacologic agents for each patient based on solid molecular evidence of efficacy.

## **ACKNOWLEDGEMENTS**

The authors wish to thank Diane Peckham, Jeffrey Lin, and Adam Asare for their assistance in the preparation of this manuscript, and Dr. Sally H. Cross and her colleagues from the United Kingdom Genome Mapping Project Centre for providing the CGI library.

## References

1. Jones, P. A. and Laird, P. W. Cancer epigenetics comes of age. *Nat.Genet.* 21: 163-167, 1999.
2. Baylin, S. B. and Herman, J. G. DNA hypermethylation in tumorigenesis: Epigenetics joins genetics. *Trends Genet.* 16: 168-174, 2000.
3. Huang, T. H.-M., Perry, M. R., and Laux, D. E. Methylation profiling of CpG islands in human breast cancer cells. *Hum. Mol. Genet.* 8: 459-470, 1999.
4. Yan, P. S., Perry, M. R., Laux, D. E., Asare, A., Caldwell, C. W., and Huang, T. H.-M. CpG island arrays: An application toward deciphering epigenetic signatures of breast cancer. *Clin. Cancer Res.* 6: 1432-1438, 2000.
5. Ahluwalia, A., Yan, P. S., Bigsby, R. M., Hurteau, J. A., Jung, S. H., Huang, T. H.-M., and Nephew, K. P. Analysis of CpG island hypermethylation in human ovarian cancer using differential methylation hybridization. *Gyn. Onc.* (in press).
6. Cross, S. H., Charlton, J. A., Nan, X., and Bird, A. P. Purification of CpG islands using a methylated DNA binding column. *Nat. Genet.* 6: 236-244, 1994.
7. Baylin, S. B., Esteller, M., Rountree, M. R., Bachman, K. E., Schuebel, K., and Herman, J. G. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum. Mol. Genet.* 10: 687-692, 2001.
8. Ehrlich, M. DNA hypomethylation and cancer. *In: M. Ehrlich (ed.) DNA Alterations in Cancer*, pp. 273-291. Natick, MA: Eaton Publishing, 2000.
9. Costello, J. F., Fruhwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., Wright, F. A., Feramisco, J. D., Peltomaki, P., Lang, J. C., Schuller, D. E., Yu, L., Bloomfield, C. D., Caligiuri, M. A., Yates, A., Nishikawa, R., Su Huang, H., Petrelli, N. J.,

- Zhang, X., O'Dorisio, M. S., Held, W. A., Cavenee, W. K., and Plass, C. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat. Genet.* 24: 132-138, 2000.
10. Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA.* 95: 14863-14868, 1998.
11. Huber, R., Hansen, R. S., Strazzullo, M., Pengue, G., Mazzarella, R., D'Urso, M., Schlessinger, D., Pilia, G., Gartler, S. M., and D'Esposito, M. DNA methylation in transcriptional repression of two differentially expressed X-linked genes, GPC3 and SYBL1. *Proc. Natl. Acad. Sci. USA.* 96: 616-621, 1999.
12. Veugelers, M., Cat, B. D., Muyldermans, S. Y., Reekmans, G., Delande, N., Frints, S., Legius, E., Fryns, J. P., Schrander-Stumpel, C., Weidle, B., Magdalena, N., and David, G. Mutational analysis of the GPC3/GPC4 glypican gene cluster on Xq26 in patients with Simpson-Golabi-Behmel syndrome: identification of loss-of-function mutations in the GPC3 gene. *Hum. Mol. Genet.* 9: 1321-1328, 2000.
13. Lin, H., Huber, R., Schlessinger, D., and Morin, P. J. Frequent Silencing of the GPC3 Gene in Ovarian Cancer Cell Lines. *Cancer Res.* 59: 807-810, 1999.
14. Santini, V., Kantarjian, H. M., and Issa, J. P. Changes in DNA methylation in neoplasia: Pathophysiology and therapeutic implications. *Ann. Intl. Med.* 134: 573-586, 2001.
15. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., and Staudt, L. M. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403: 503-511, 2000.

16. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286: 531-537, 1999.
17. Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., Sampas, N., Dougherty, E., Wang, E., Marincola, F., Gooden, C., Lueders, J., Glatfelter, A., Pollock, P., Carpten, J., Gillanders, E., Leja, D., Dietrich, K., Beaudry, C., Berens, M., Alberts, D., and Sondak, V. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406: 536-540, 2000.
18. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A. L., Brown, P. O., and Botstein, D. Molecular portraits of human breast tumours. *Nature* 406: 747-752, 2000.

## Figure Legends

Fig. 1. Representative results of differential methylation hybridization. *A*, Breast tumor and normal amplicons were prepared as described in the text, fluorescently-labeled with Cy5 and Cy3, respectively, and co-hybridized to microarray slides containing 7,776 CpG island tags. The hybridization output is the measured intensities of the two fluorescence reporters false-colored with red (tumor) and green (normal) and overlaid with each other. “Yellow” spots indicate equal amounts of bound DNA from each amplicon, signifying no methylation differences between tumor and normal genomes. Spots hybridized predominantly with tumor amplicon, but not with normal amplicon, would appear red and are indicative of hypermethylated CpG island loci, present in the tumor genome. The expanded view of the box area is shown in panels *B*, *C*, and *D*. The SC76F1 locus (marked by arrows), a CpG island of the *glypican 3* gene, appears to be hypermethylated in the breast tumor of patient 241, but not in patients 151, and 109. Green spots, seen in the breast tumor of patient 151, denote hypomethylation of normally methylated repeat sequences.

Fig. 2. Representative results of methylation analysis by Southern hybridization. Breast tumor and normal amplicons (~1 µg each) were prepared as described in the text and subjected to Southern analysis. T, tumor amplicon; N, normal amplicon. The nylon membranes were hybridized with <sup>32</sup>P-labeled fragments derived from CpG island loci as indicated on the left. The presence of a hybridized DNA fragment in the lane suggests that a test locus was protected from methylation-sensitive restriction by both *Hpa*II and *Bst*UI and was amplified by linker-PCR (see Materials and Methods). As indicated, a locus that does not contain the *Hpa*II and *Bst*UI recognition sites was used as control to check DNA loading in each lane.



Fig. 3. Methylation profiles for breast tumors. *A*, Hierarchical clustering of methylation data. The dendrogram at the *top* lists the 17 breast tumors studied and indicates the degrees of relatedness between tumors. The row corresponds to each of 496 CpG island loci selected for methylation analysis. CpG island loci left blank are those with the normalized Cy5/Cy3 ratios <1.5; these loci exhibit no methylation differences between normal and tumor DNA, or have no significant hybridization signals. CpG islands colored red (the normalized Cy5/Cy3 ratios  $\geq 1.5$ ) are those with DNA hypermethylation in tumor DNA. Three subclusters indicated by black bars are described in greater details in panel *C*. *B*, Multidimensional-scaling plot of the 17 breast tumors analyzed. In this three-dimensional plot, tumor samples with similar methylation patterns lie closer to each other than those with dissimilar profiles and are marked with the same color. As shown by dotted lines, 15 of the 17 tumors are grouped into a hormone receptor-positive (red) and two hormone receptor-negative (blue and green) clusters. *C*, Subclusters of CpG island loci segregating breast tumors according to their hormone-receptor status. The color-coded dendrogram at the top indicates the degrees of relatedness among the 17 tumors analyzed. Nuclear grade and the hormone-receptor status are listed directly under each patient analyzed. ND: sequence not determined. A CpG island locus (marked by an arrow) for the *glypican 3* gene was studied in greater detail (see Fig. 4).

Fig 4. *A*. Methylation analysis of the *glypican 3* (*GPC3*) gene promoter in breast tumors. The positions of the methylation-sensitive *Bst*UI sites and the probe in the *GPC3* region are indicated (nucleotide positions 98707-99471, GenBank accession number AF003529). \**Bst*UI: This site is methylated in the inactive X chromosome. Breast tumor DNA (10  $\mu$ g) was treated with *Mse*I and *Bst*UI and subjected to Southern analysis. C: Control DNA digested with *Mse*I only; N: Normal

DNA digested with *Mse*I and *Bst*UI. DMH analysis was conducted in breast tumors that are underlined. HM, hypermethylated DNA fragments (major bands: 770 and 570-bp); NM, normally methylated DNA fragment (490-bp) in the inactive X chromosome; UM, unmethylated DNA fragments (360 and 130-bp). The percentage of hypermethylation was calculated as the total intensities of HM relative to the combined intensities of all the bands (HM+NM+UM). *B*, Association of *GPC3* hypermethylation and the hormone receptor status of breast tumors. High levels of *GPC3* methylation are significantly associated with the hormone receptor-negative tumors ( $P=0.005$ ). *C*, Methylation analysis of the *glypican 3* (*GPC3*) gene promoter in 6 breast cancer cell lines. Genomic DNA (10  $\mu$ g) was treated with *Mse*I and *Bst*UI and subjected to Southern analysis. Extensive hypermethylation of the *GPC3* CpG island was observed in 5 (except T47-D) of these 6 breast cancer cell lines examined. *D*, RNA analysis of the *GPC3* gene in normal breast epithelial cells (N), and untreated (U) and 0.75  $\mu$ M 5-aza-2'-deoxycytidine treated (6 days) breast cancer cells (T). Northern blot analysis was conducted for cell lines T47-D, ZR-75, and Hs578t, whereas semiquantitative RT-PCR analysis was conducted on the MDA-MB-231 cells. The levels of *GPC3* expression were normalized with that of  $\beta$ -actin. *GPC3* was expressed in normal breast epithelial cells but was markedly down-regulated or absent in methylated breast cancer cell lines, ZR-75, Hs578t, and MDA-MB-231. Expression of *GPC3* was partially restored in these cell lines after the demethylation treatment.

Table 1. Verification of DMH results by Southern blot analysis.

Patient	CpG island loci									
	SC13B2		SC76F1		CpG15E5		CpG6H12		CpG5H12	
	<sup>*</sup> Southern	<sup>†</sup> DMH	Southern	DMH	Southern	DMH	Southern	DMH	Southern	DMH
25	+	3.0	+	2.0	+	5.9	-	1.6	+	2.4
31	+	4.1	+	1.8	+	10.8	+	2.2	+	5.6
43	(-	1.5) <sup>‡</sup>	-	1.4	-	ND	-	0.7	(-	3.8)
45	-	1.4	+	5.1	-	ND	-	1.0	+	5.2
65	+	3.0	+	1.8	+	1.8	ND	1.1	+	2.2
83	+	2.1	-	1.2	+	2.1	+	3.3	+	3.2
91	(-	1.5)	-	1.3	+	4.6	+	3.0	+	3.7
109	-	1.1	-	1.2	-	1.1	-	0.9	-	1.0
119	+	1.7	+	1.7	+	ND	-	0.9	(+	1.2)
129	+	1.7	-	1.3	-	1.4	-	1.3	+	1.6
151	(+	1.4)	-	1.1	+	1.8	-	1.2	+	1.9
155	+	4.2	+	1.8	+	7.2	+	2.7	+	7.8
157	-	0.9	-	1.2	+	1.6	-	1.3	(+	1.1)
159	+	4.0	+	2.3	+	3.8	+	2.2	+	4.7
167	+	ND <sup>§</sup>	+	1.6	+	1.9	-	0.8	+	2.0
241	+	2.5	+	4.6	+	9.8	+	1.7	+	6.3
243	+	1.8	+	2.0	+	2.8	-	1.2	+	4.0

\*Southern blot visual scores: "+" indicates hypermethylation in tumor DNA; "-" indicates no methylation differences between tumor and normal DNA. See Fig. 2 for selected examples.

<sup>†</sup>Adjusted DMH Cy5/Cy3 ratios. As described in the text, the cutoff ratio for hypermethylation in tumor DNA is set at  $\geq 1.5$ .

<sup>‡</sup>Loci with ratios enclosed in a bracket depict those not confirmable by Southern analysis.

<sup>§</sup>ND: data not determined

Fig. 1 Yan *et al.*

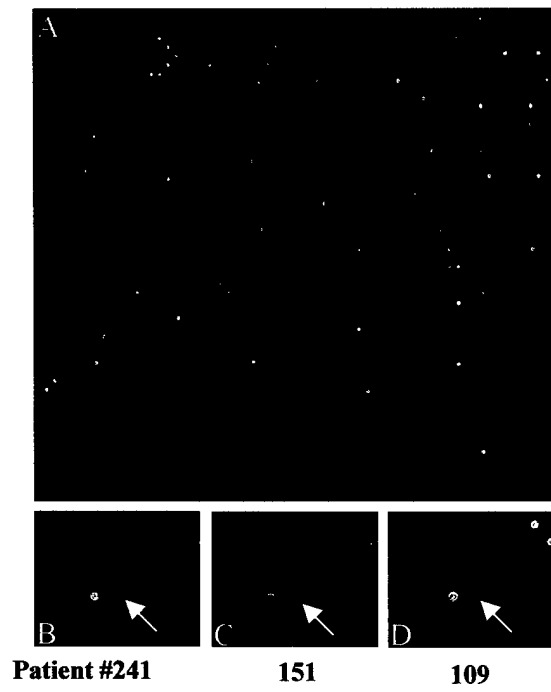


Fig. 2 Yan *et al.*

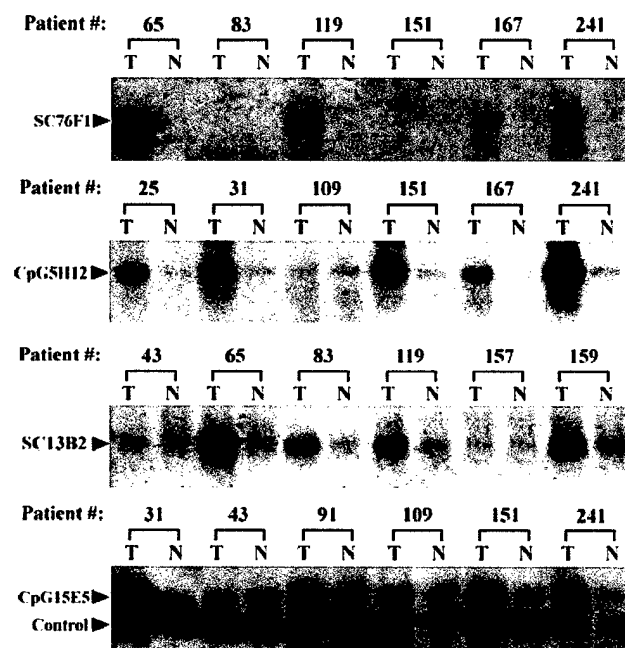


Fig. 3 Yan *et al.*

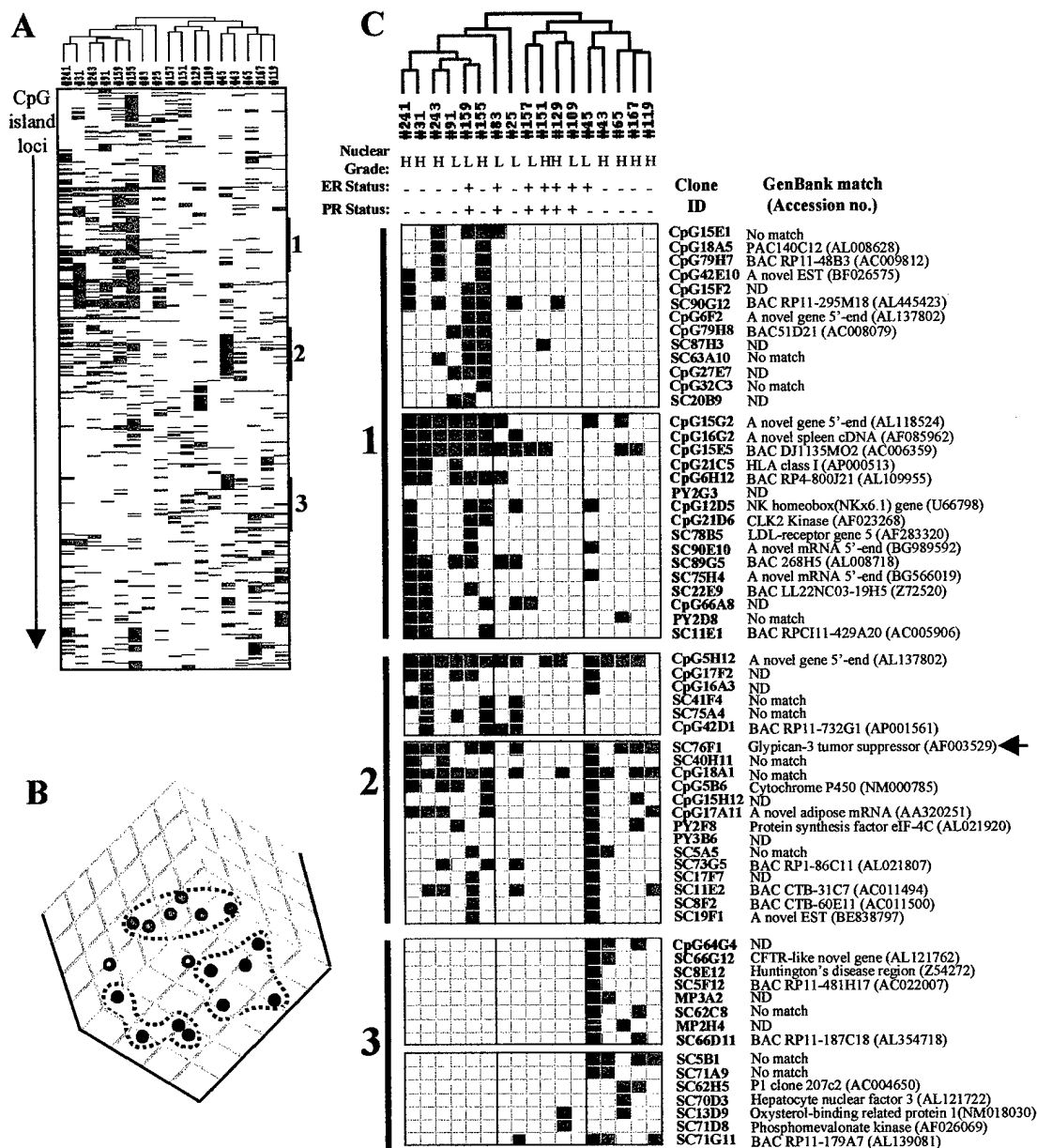


Fig. 4 Yan *et al.*

